

**COMPREHENSIVE EVALUATION SYSTEM ON BAHASA INDONESIA
LEARNING OUTCOME****Wagiran[✉], Dandan Supratman, Fathur Rokhman**

Postgraduate Program of Semarang State University, Indonesia

Article Info*Article History:*

Accepted October 2015

Approved October 2015

Published November

2015

*Keywords:**Measurement, Assessment,**Evaluation System,**Comprehensive, Learning**Outcomes, Indonesian***Abstract**

This study aims to develop a comprehensive evaluation system on Bahasa Indonesia learning outcome for Junior High School students based on Gall & Borg model. The study has some findings. First, the SWOT analysis results show the need of developing a comprehensive evaluation system on Indonesian learning outcomes at grade level. Second, the rules of the comprehensive evaluation system are packed into three parts; the conceptual rules of learning outcomes evaluation, the implementation rules of the evaluation of learning outcomes at the beginning, on the process, and at the end of the study, and follow-up of the learning outcomes evaluation. Third, the research product in the form of a comprehensive evaluation system of learning outcomes manual book for Indonesian teachers is considered feasible by expert and very well perceived by the teacher. Fourth, the try-out results with the technique of one-sample t-test (one-tail) at a 95% confidence level showed a significant increase on teachers' pedagogical competence. This is shown by the group A, t-value 6.6149 is greater than the table of t-values 1746, group B, t-value 7.8257 is greater than the table of t-values 1729, and the group C 10.0857 t-count is greater than the table of t-values 1697. Thus, this comprehensive Indonesian learning outcomes evaluation system is effective to improve teachers' pedagogical competence.

© 2015 Semarang State University

[✉] Correspondence Address :

Unnes Bendan Ngisor Campus, Semarang, 50233

E-mail: pps@unnes.ac.id

ISSN 2085-4943

INTRODUCTION

This research topic is based on the fact that there is a gap between the practice of Indonesian learning outcomes evaluation junior high schools with the learning outcomes evaluation theory that should be the reference. The gap occurs because some theories of language learning evaluation are ignored in practice or because some evaluation theories are conceptually ideal but difficult to apply in practice. A comprehensive study on this matter is needed to establish a harmony between the theory and practice of Indonesian learning outcomes evaluation.

Some researches on learning outcomes evaluation indicate that there is a number of malpractice in conducting the evaluation. Purwati and Wulandari (2008), Sulistya (2009), and Wagiran and Doyin (2010) reveal malpractice in the evaluation, and the findings is thought as only a fraction of the actual reality. This condition is like the "iceberg" phenomenon.

The malpractice on learning outcomes evaluation is not the main thing that we need to be worried about, it is its negative impact that needs our attention. Evaluation malpractice can cause a negative washback on the way teachers teach and how students learn and can obscure the actual achievement of educational goals. This has been demonstrated by the results of the evaluation washback fit on several international research journals conducted by Alderson and Wall (1993), Biggs (1995, 1996), and Burrows (2004).

Alderson & Wall (1993) show that improper evaluation affects the way teachers teach and the way students learn. They also mention that tests are strong determinant, both positively and negatively, on what have happened in the classroom. Biggs (1995, 1996) proves that tests definitely affect teaching practices. He explains that tests can foster the teaching achievements. Burrows (2004) in his research *Washback in Classroom-Based Assessment: A Study of the Washback Effect in the Australian Adult Migrant English Program 113* shows that teachers will tend to put the real communicative

competence aside if the assessment is not based on communicative approach. They will also ignore classroom assessment if it has no significant contribution on students' learning achievement, and if the skills being assessed are not in line with the goals of the learning, it will make the learning unmeaningful.

Based on the description of the research backgrounds and relevant researches, a comprehensive learning outcome evaluation system on Bahasa Indonesia needs to be developed as an effort to reduce the negative impact on Indonesian learning outcomes. In connection with that, the purpose of this study is as follows: (1) Describing the need for a comprehensive Indonesian learning outcomes evaluation system on junior high school at grade level, (2) Formulating a comprehensive Indonesian learning outcomes evaluation system rules in junior high school at grade level, (3) Designing a prototype of a comprehensive evaluation system of learning outcomes manual book for Indonesian teachers, (4) Proving the effectiveness of the comprehensive evaluation system of learning outcomes manual book for Indonesian teachers in improving the teachers' pedagogical competence.

RESEARCH METHOD

To develop a comprehensive Indonesian evaluation system of learning outcomes manual book for Indonesian teachers, Research and Development model by Gall and Borg is used. In this research, the ten stages proposed by the model are simplified into four stages, in accordance with the number of research purposes. The four stages are: (1) the need analysis of the product, (2) the rules formulation, (3) product development, and (4) validation and dissemination of the product.

The teacher needs on a comprehensive Indonesian evaluation system of learning outcomes are gathered using SWOT analysis. The collection of data related to the weaknesses and strength of learning outcomes evaluation system was done by using in-depth interviews, focus group discussions, and document analysis.

The collection of data relating to the rules of the comprehensive Indonesian evaluation system of learning outcomes done by doing literature review with the results of SWOT analysis. The product development of a comprehensive Indonesian evaluation system of learning outcomes manual book for Indonesian teachers was carried out through an expert assessment and a series of try-outs. There are four experts taking part in this research: two experts on evaluation, one expert on teaching and learning, and one expert on non-lesson book. Limited scale trial was conducted on a group of Indonesian teachers in Semarang. Wide scale trial was conducted on three groups of Indonesian teachers in the district of Magelang.

The primary data was gathered by doing interviews on thirty-five Indonesian teachers in public and private junior high schools in Central Java. The secondary data was the document of Indonesian learning outcomes evaluation used by the Indonesian teachers in junior high school teacher in Central Java.

Data on the score of the test quality were analyzed in percentage, and the data on the increasing score of the teachers' competence were analyzed using nonparametric statistics. Data in the form of verbal description were analyzed qualitatively in the following steps: (1) gathering the data, (2) organizing the data, (3) grouping the data, (4) synthesizing the data to find the patterns, important data which is in line with the goal of the research, to be a substantive theory, and (5) deciding which can be reported as the result of the research (Bogdan & Biklen 2003).

RESULTS AND DISCUSSION

There are two main research findings, the need of a comprehensive Indonesian learning outcomes evaluation system and its rules. The following description is on the findings.

Need of a Comprehensive Indonesian Learning Outcomes Evaluation System

To get an idea of the need of a comprehensive Indonesian learning outcomes

evaluation system, a SWOT analysis was performed. Based on the results of the evaluation document analysis, in-depth interviews with Indonesian teachers as the executor of the evaluation of learning outcomes, as well as the study of literature, the SWOT analysis shows the following findings.

The determinants of successful implementation of a comprehensive Indonesian learning outcomes evaluation system is professional Indonesian teachers. Professional teacher has full authority to plan and develop an assessment system of learning outcomes, and use it for the sake of learning at grade level (Permendiknas No. 20/2007 and Permendikbud No. 66/2013). Professional teacher will be able to utilize its authority to take advantage of development opportunities, reduce the weaknesses, and prevent threats that could damage the system of learning outcomes evaluation.

SWOT analysis results show that the system of Indonesian learning outcomes evaluation has not been managed appropriately. Instead of improving the quality of the teaching learning process and its outcomes, evaluation activities obtained precisely the opposite. Evaluation system that is not managed correctly resulted in the Indonesian learning process that is no longer based on communicative competence as the demands of the curriculum but rather emphasizes that pragmatic to pursue UN values. Mannered character language use as a result of learning accompaniment is replaced by instant thinking patterns in achieving short-term goals. In terms of learning achievement, there has also been a reduction of the learning objectives of the three aspects (attitudes, knowledge and skills) that should be reflected in the communicative competence. Only knowledge is achieved. The Indonesian score no longer reflects the actual language competence.

The condition is in accordance with the evaluation wash back theory that has been proven by the Wall and Alderson (1993), Shohamy, Schmidt, and Ferman (1996), Messick (1996), McNamara (2001), and Saif (2006).

Table 1. SWOT Analysis on the need of a comprehensive Indonesian learning outcomes evaluation system

Internal Condition	
Strength	Weakness
<p>The Indonesian teachers have fulfilled the requirements of professional teachers;</p> <p>The planner and the doer of the evaluation of learning outcomes in the classroom are the teachers;</p> <p>Daily test has high percentage in deciding the students' learning outcomes;</p> <p>The teachers as the evaluator know the students' condition and characteristics.</p>	<p>Affective assessment is ignored as it is considered has little contribution in students' learning achievement;</p> <p>Assessment on psychomotor aspect is ignored as it takes a lot of time;</p> <p>There are evaluation instruments which do not fulfil the criteria resulting invalid assessment result;</p> <p>Evaluation follow up has not been conducted professionally</p>
External Condition	
Opportunity	Threat
<p>National education goal demands a comprehensive achievement on cognitif, affective, and psychomotor aspects (UU Sisdiknas);</p> <p>Nasional Education department plan demands an evaluation system in accordance with the future demand (Permendiknas No, 2, 2010);</p> <p>Teachers' pedagogical competence demands the teachers do assessment and evaluation on the teaching and learning process and the learning outcomes (Permendiknas No 16/2007);</p> <p>Modern language learning theory demans language evaluation communicatively (Clark, 1987; Brindley, 1997)</p>	<p>National Exam negative washback limits the evaluation process on grade level;</p> <p>A policy to conduct an integrated daily test makes the daily test function minimaly;</p> <p>Parents' and officials' pride on cognitive aspect make affective and psychomotor aspects look useless;</p> <p>Learning outcomes evaluation is still unable to increade the teaching learning process and its outcomes;</p> <p>Indonesian score does not imply the students' real ability in using Bahasa Indonesia</p>

Wall and Alderson (1993) in an article entitled *Examining Washback: The Sri Lankan Impact Study* in the journal *Language Testing*, Vol. 10, No. 1, 41-69 reveals that the language exam often cause adverse effects in language teaching. This condition occurs in Indonesian language learning today that received a negative effect on the implementation of the UN.

Shohamy, Schmidt, and Ferman (1996) in *Language Testing*, Vol. 13, No. 3 modify national test conducted over two consecutive years of two languages, namely Arabic and English. The modification can create a positive impact on teaching activities despite the fact that the negative behaviors of the tests still remain. They believe that such tests create a meaningful change

and a force sufficient to change without a need to provide new training and curriculum. The conditions as described by Shohamy, Schmidt, and Ferman is expected to occur as a result of the development of a comprehensive Indonesian learning outcomes evaluation system.

Messick (1996) in an article entitled *Validity and Washback* in the journal *Language Testing* Vol. 13, No. 3 proves that authentic assessment and assessment has a direct positive effect on the quality improvement of learning process and learning outcomes. Saif (2006) in an article entitled *Aiming for positive washback: a case study of international teaching assistants*, published in the journal *Language Testing*, Vol. 23, No. 1, shows the possibility of

the creation of a positive effect by focusing on the background of the exam development process and anticipate the conditions that enable the emergence of a positive effect. Saif (2006) further reported the results of empirical research shows the influence of multi-phase testing on a need-based test. The positive effects of evaluation as found Messick and Saifare expected to arise from the development of a comprehensive Indonesian learning outcomes evaluation system because this system accommodates authentic assessment in the evaluation include evaluation of communicative language that is natural and direct.

McNamara (2001) in the journal *Language Testing*, Vol. 18, No. 4 wrote an article about the *Language Assessment as Social Practice: Challenges for Research*. In that article, McNamara recommends a rethinking of language assessment in accordance with the progress of language learning theory-driven post modernization. This evaluation system is a

follow up of the recommendations applied in the Indonesian evaluation.

SWOT analysis as well as its compatibility with some of the recommendations reinforce the need of redefining the rules of reinforcement will need to be redefined norm of a comprehensive Indonesian learning outcomes evaluation system as a guide for Indonesian teachers to "carry out the assessment and evaluation on learning processes and learning outcomes" and "utilizes the result of evaluation to improve the learning quality" (PermendiknasNo. 16 of 2007).

The Rules of comprehensive Indonesian learning outcomes evaluation system

The rules of a comprehensive Indonesian learning outcomes evaluation system can be divided into three main rules: (1) conceptual rules, (2) operational rules, and (3) follow-up rules. These three main rules are described into six-section manual guide book for teachers, as the following Figure 1.

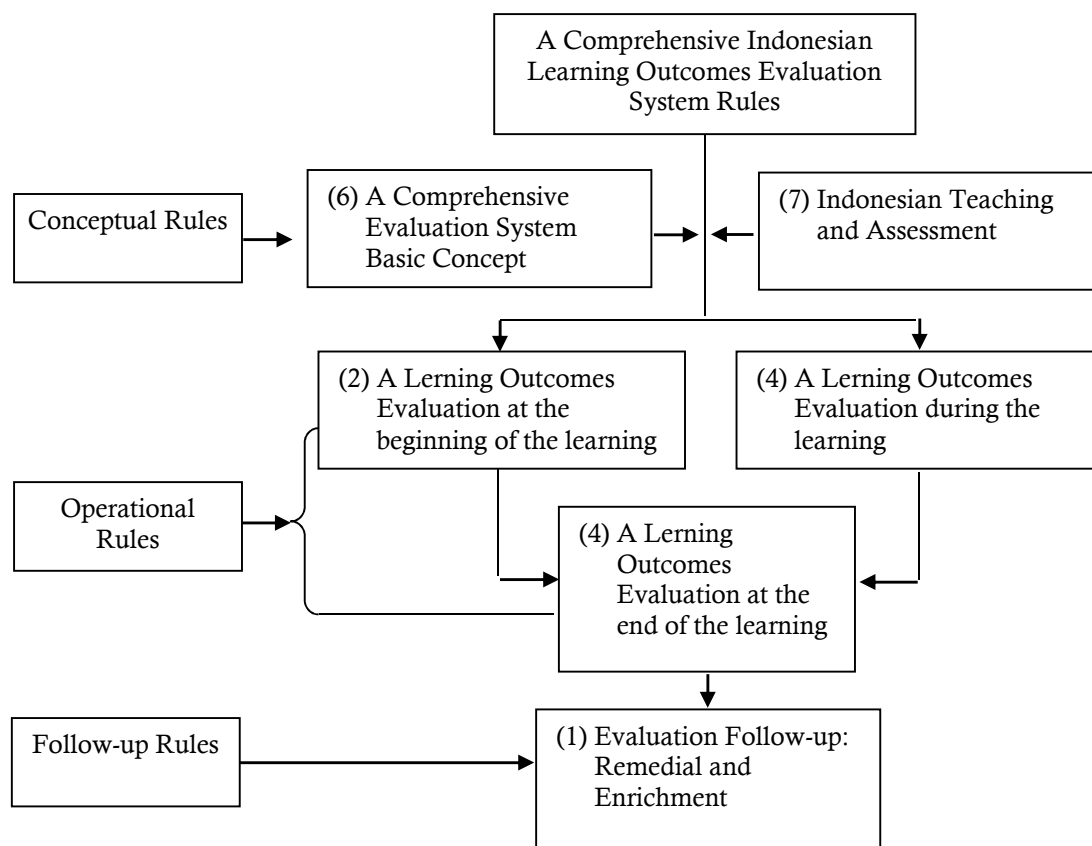


Figure 1. A Comprehensive Indonesian Learning Outcomes Evaluation System

Related to the conceptual rules, two things were examined: (1) the basic concept of a comprehensive evaluation system and (2) Indonesian learning and assessment. The basic concept of a comprehensive evaluation system is presented in Chapter I, which consists of three sections: the basic concepts of evaluation, a comprehensive evaluation system, and the ethics of a comprehensive evaluation. Indonesian learning and assessment embodied in Chapter II, which contains the characteristics of Indonesian learning and Indonesian assessment.

Related to the operational rules, three things were examined, the results of a comprehensive evaluation system of Indonesian learning at the early stage of learning is presented in Chapter III, the evaluation during the learning process is presented in Chapter IV, and the evaluation at the end of study is presented in Chapter V. Each chapter contains an evaluation target and the stages of implementation in accordance with the evaluation of the learning phase.

Related to the follow-up rules to the comprehensive evaluation of learning outcomes, three things were examined: remedial teaching, learning enrichment, as well as remedial and enrichment of specific cases. The rules is presented in Chapter VI.

A Comprehensive Indonesian Learning Outcomes Evaluation System Development

The development of the comprehensive Indonesian learning outcomes evaluation system passed three stages: the development of the manual book for teachers prototype, expert validation and the first revision, trying out the manual book on a limited scale and the second revision, and effectiveness testing on a wide scale.

The development of the manual book for teachers prototype is based on the formulation of a comprehensive evaluation system rules and the rules of preparation of learning outcomes non-lesson book to guide teachers. The pattern of the prototype development is illustrated in the following Figure 2.

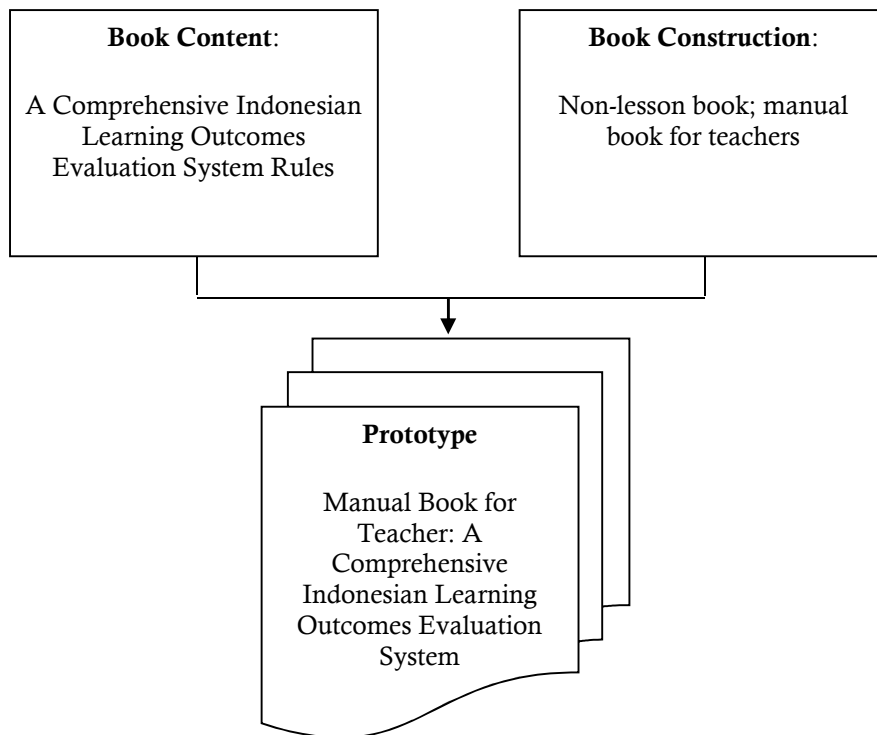


Figure 2 . Prototype Manual Book for Teacher: A Comprehensive Indonesian Learning Outcomes Evaluation System

TITLE			
PREFACE		ii	} The Beginning Part
TABLE OF CONTENT		iv	
TABLE OF TABEL		vii	
Chapter I	A Comprehensive Learning Outcomes Evaluation System	1	} The Main Part
Chapter II	Indonesian Learning and Assessment Characteristics	47	
Chapter III	Implementation of A Indonesian Comprehensive Learning Outcomes Evaluation System at the beginning of the lesson	134	
Chapter IV	Implementation of A Indonesian Comprehensive Learning Outcomes Evaluation System during the lesson	151	
Chapter V	Implementation of A Indonesian Comprehensive Learning Outcomes Evaluation System at the end of the lesson	161	
Chapter VI	Follow-up of the Indonesian Comprehensive Learning Outcomes Evaluation System	183	
BIBLIOGRAPHY		196	} The End Part
INDEX		204	
GLOSSARY		206	

Figure 3. Structure Book Content

Based on the patterns of development, the manual book prototype contains a comprehensive Indonesian learning outcomes evaluation system. After being evaluated by some experts of assessment, language learning, and non-lesson book, the prototype follows the following structure Figure 3.

Manual book for teachers prototype: A Comprehensive Indonesian Learning Outcomes Evaluation System was evaluated by four experts: Prof. Dr. Burhan Nurgiyantoro, M.Pd. (UNY) and Sumarwati, M.Pd. (UNS), as assessment expert, Prof. Dr. Sukirno, M.Pd. (UMP) and Prof. Dr. Suherli Kusmana, M.Pd. (Puskurbuk) as non-lesson books experts. The expert assessment results are presented as follows.

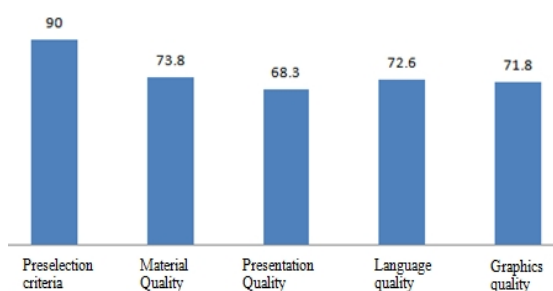


Figure 4. Experts' assessment on Manual Book for teacher Prototype.

Comprehensive Learning Outcomes Evaluation System

Based on 20 non-lesson book pre selection criteria, only one criterion that has not fulfilled the requirements: that "It has been published in the last 3 years, and has ISBN numbers. The advice given is to publish the book at a bonafide book publisher. The suggestions have been implemented by having the book published by *Mediatama Surakarta*, entitled Manual book for teachers: publishing the publication educator guide: A Comprehensive Indonesian Learning Outcomes Evaluation System with ISBN 978-602-449-439 and is currently listed in the collection of non-lesson book in Puskurbuk.

Based on fifteen criteria of non-lesson book (manual book for teachers), there are three indicators considered sufficient. They are (1) a description of material still poses multiple interpretations of the reader and have been revised by rearranging the idea carefully, (2) the substance of the Chapter IV has not given an understanding of the objectives and stages of the evaluation process of learning the Indonesian language, has been improved by describing the material focused on the assessment of learning

outcomes in the learning process with the mechanism of formative evaluation, and (3) the substance of the Chapter V has not given an understanding of the objectives and stages of evaluation of learning outcomes Indonesian repaired by description of the material focused on the achievement of learning outcomes with mechanisms summative evaluation .

Of the ten criteria for the presentation of the non-lesson book, there are four indicators considered sufficient with some improvements. (1) the material presentation that has not developed reader beliefs about religious attitudes has been improved by adding the sentences that lead to a sense of gratitude to God. (2) The presentation of the material that has not been encouraging readers to develop social attitudes has been improved by adding words that evoke social attitudes. (3) The presentation of the material that has not yet developed the knowledge has been improved by providing concrete examples of the application of knowledge to the problems of evaluation. (4) Presentation of the material that has not been able to develop the skills to evaluate learning outcomes has been improved by providing examples of the implementation of evaluation at an early stage, during the process, and at the end of the lesson.

Of the five criteria for good language usage, there are two indicators were considered sufficient with some improvements. (1) The language use are considered not having language beauty and has been improved rearranging sentences and paragraphs that meet the elements of beauty. (2) The words, that are not written based on the rule of *EYD*, have been improved by reediting the manuscript.

Based on ten criteria for graph assessment, the overall assessment in the sufficient category with some suggestions for improvement: (1) use meaningful illustration, (2) reset the cover, (3) add book content on the back cover pages of the back cover, (4) put author information in the book, (5) present the content of the book using concept map, (6) choose appropriate font, for example, Calibri, and (7) add a glossary. The suggestions have been

followed up by publishers by relayouting the book.

A Comprehensive Indonesian Learning Outcomes Evaluation System Try-outs

Limited scale try-out was conducted on a group of Indonesian teachers in Semarang. The trial results showed that the manual book for teachers: *A Comprehensive Indonesian Learning Outcomes Evaluation System* can be accepted by Indonesian teachers with the following results. (1) The book is very helpful for teachers in conducting the assessment and evaluation of learning outcomes (65%), (2) the book is useful to guide conduct an evaluation of Indonesian learning outcomes (73%), (3) the evaluation system can be implemented easily (85 %), (4) the book is easy to understand (55%), and (5) the book has interesting presentation (50%).

On the large-scale trials, the trials which were conducted in Magelang District MGMP in three periods of the education training shows the following results. The results of trials to test the one-sample t-test (one-tail) at a 95% confidence level showed no significant increase in competence of "carrying out the assessment and evaluation of processes and learning outcomes". This is shown by the group A, t-value 6.6149 is greater than t-table 1746, group B, t-value 7.8257 is greater than t-table 1729, and the group C 10.0857 t-count is greater than t-table 1697. Thus, the comprehensive Indonesian learning outcomes evaluation system is effective to improve Indonesian teachers' pedagogical competence, especially in "carrying out the assessment and evaluation of processes and learning outcomes" and "utilizing the results of the evaluation to improve learning".

CONCLUSION

SWOT analysis show the need for a comprehensive Indonesian learning outcomes evaluation system at grade level.

The comprehensive Indonesian learning outcomes evaluation system needs the following rules: (1) the rules of conceptual comprehensive Indonesian learning outcomes evaluation system

consisting of the basic concepts of evaluation, a comprehensive evaluation system, and the ethics of a comprehensive evaluation, (2) Characteristics of learning, aspects of assessment and learning outcomes assessment techniques Indonesian, (3) the rules in conducting the evaluation of learning outcomes at the beginning, during the process, and at the end of the lesson, and (4) follow-up rules.

The research product in the form of manual book for teachers: a Comprehensive Indonesian Learning Outcomes Evaluation System is considered feasible by experts (evaluation experts, language learning experts, and the non-lesson book expert. In addition, the book is very well perceived by the teacher.

The trial results with the technique of one-sample t-test (one-tail) at a 95% confidence level showed no increase pedagogical competence of teachers is significant. This is shown by the group A, t-value 6.6149 is greater than t-table 1746, group B, t-value 7.8257 is greater than t-table 1729, and the group C 10.0857 t-count is greater than t-table 1697. Thus, a comprehensive Indonesian learning outcomes evaluation system is effective to improve teachers' pedagogical competence.

The comprehensive Indonesian learning outcomes evaluation system is implemented as the main material for education and training program evaluation of learning outcomes for Indonesian teachers.

The results of a comprehensive Indonesian learning outcomes evaluation system at grade level get a decent portion in determining learning outcomes.

The results of this study becomes a reflection for the government to think of changing the policy of the National Examination into National Evaluation.

REFERENCES

- Akhadiah, Sabarti. (1988). *Evaluasi dalam Pengajaran Bahasa*. Jakarta: Proyek Pengembangan Lembaga Pendidikan Tenaga Kependidikan Dirjen. Dikti.
- Alderson, J. C., & Wall, D. 1993. Does Washback Exist? *Applied Linguistics*, 14, 115–129.
- Biggs, J. B. (1995). Assumptions underlying new approaches to educational assessment. *Curriculum Forum*, 4(2), 1–22.
- Biggs, J. B. (Ed.). (1996). *Testing: To educate or to select? Education in Hong Kong at the cross-roads*. Hong Kong: Hong Kong Educational Publishing.
- Bogdan, R. C & Biklen, S. K. (2003). *Qualitative Research for Education: An introduction to Theories and. Methods* (4th ed.). New York: Pearson Education group
- Brindley, G. (1997). Assessment and reporting in second language programs: Purposes, problems and pitfalls. In E. S. L. Li & G. James (Eds.) *Testing and evaluation in second language education*. Hong Kong: Language Centre, Hong Kong University of Science and Technology.
- Burrows, C. (2004). Searching for washback: The impact of assessment in the Certificate in Spoken and Written English. In G. Brindley & C. Burrows (Eds.), *Studies in immigrant English language assessment: Vol. 2*. Sydney, Australia: National Center for English Language Teaching and Research.
- Clark, J. L. (1987). "Classroom assessment in a communicative approach". *British Journal of Language Teaching*, 25(1), 9-19.
- Djiwandono, M. Soenardi. 1996. *Tes Bahasa Dalam Pengajaran*. Bandung: ITB.
- Doyin, Mukh dan Wagiran. 2007. "Pengembangan Penilaian Bahasa dan Sastra Indonesia Berbasis Komunikatif di SMP". Laporan Penelitian Universitas Negeri Semarang.
- Gronlund, N.E., dan Linn, R.L. (1990). *Measurement and evaluation in teaching*. New York: McMillian Publishing Company.
- Masden, Harold S. 1983. *Techniques In Testing*. New York: Oxford University Press.
- McNamara (2001) dalam jurnal *Language Testing*, Vol. 18, No. 4
- McNamara, T. (2001). "Language Assessment as Social Practice: Challenges for Research" *Measuring second language performance*. London: Longman.
- Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13, 241–256
- Purwanti, Ani dan Irni Wulandari. 2008. "Studi Kualitas Soal Ujian Akhir Sekolah Berstandar Nasional (UASBN) Mata Pelajaran Matematika Provinsi DKI Jakarta Wilayah Jakarta Timur". Laporan Penelitian Universitas Terbuka.
- Saif (2006) dalam artikel yang berjudul *Aiming for positive washback: a case study of international teaching assistants* yang dimuat pada jurnal *Language Testing*, Vol. 23, No. 1

- Shohamy, E., Donitsa-Schmidt, S., & Ferman, I. (1996). Test Impact revisited: Washback effect over time. *Language Testing*, 13, 298–317.
- Stuflebeam, D.L. and Shinkfield, A.J. 1986. *Systematic Evaluation*. Boston:
- Sulistya. 2009. “Validitas Prediksi Hasil Belajar Bahasa Inggris”. *Tesis*. Universitas Negeri Yogyakarta.